A Physical Method for the Calibration of the AVHRR/3 Thermal IR Channels. Part II: An In-Orbit Comparison of the AVHRR Longwave Thermal IR Channels on board MetOp-A with IASI

JONATHAN MITTAZ AND ANDREW HARRIS
CICS/ESSIC, University of Maryland, College Park, College Park, Maryland

(Manuscript received 15 August 2010, in final form 8 April 2011)

ABSTRACT

Obtaining stable and accurate satellite radiances for climate change research requires extremely high standards for satellite calibration. Many satellite sensors do not currently meet the accuracy criteria, especially heritage sensors such as the Advanced Very High Resolution Radiometer (AVHRR), which shows scene temperature–dependent trends and biases of up to 0.5 K. Recently, however, a detailed study of the AVHRR/3 prelaunch data showed significant problems with both the calibration algorithm and the prelaunch data and indicated that the inherent accuracy of the AVHRR may actually be quite high. A new approach has been suggested that fixed many of the issues with the current (operational) calibration, but has not yet been applied to the in-orbit case. In this paper the behavior of the AVHRR in orbit is examined and compared to the operational AVHRR radiances from the Meteorological Operation (MetOp)-A with those based on the new calibration to radiances derived from the Infrared Atmosphere Sounding Instrument (IASI). It is shown that the current AVHRR calibration does indeed introduce large (0.5 K) biases, but these biases are remarkably stable. It is further shown that, with some modification related to differences between the prelaunch test environment and the in-orbit environment, a physically based AVHRR calibration can match IASI to better than ~0.05 K, which is an order of magnitude better than what is currently available. Finally, it is shown that, while the new calibration is capable of providing accurate and stable radiances for the nadir view, off-nadir biases of up to 1.5 K still exist at the largest zenith angles and at the coldest scene temperatures (~210 K). For surface temperature determination, however, the scan angle bias is very small (~0.02 K), implying that the new AVHRR calibration will provide a significant improvement to, for example, sea surface temperature measurements, one of the Global Climate Observing System (GCOS)-designated essential climate variables.

1. Introduction

In this era of climate change research there is an ever-increasing need for accurate and stable satellite data. This, however, places severe constraints on the accuracy of the instrument calibration with required biases on the brightness temperatures of <0.1 K together with a stability of <0.05 K decade\(^{-1}\) (e.g., Ohring et al. 2005). These requirements are at or even beyond what many operational sensors can achieve using current calibration methodologies, and they are particularly problematic for those sensors with a long history. For example, one of the longest continuously used sensors is the Advanced Very High Resolution Radiometer (AVHRR), versions of which have flown from the late 1970s, but whose original design specifications were 1 K for the early sensors and 0.5 K for the later AVHRR/3 series. If the AVHRR sensors can do no better than this, then their use for climate change studies will be questionable.

Recent initiatives in cross-calibrating different sensors are now shedding light on the relative accuracy of different sensors. Projects such as the Global Space-Based Inter-Calibration System (GSICS; online at http://www.star.nesdis.noaa.gov/smcd/SPB/calibration/icvs/GSICS/index.php) now provide regular intercomparisons and are highlighting some of the issues with instrument calibration. One of the baseline instruments for GSICS is the Infrared Atmospheric Sounding Interferometer (IASI), a hyperspectral sounder that is thought to be well calibrated (Blumstein et al. 2007). IASI is, in many ways, an ideal instrument for studying the AVHRR calibration because the IASI and AVHRR instruments are both on the Meteorological Operation (MetOp)-A platform, and

Corresponding author address: Jonathan Mittaz, Suite 4001, 5825 University Research Court, CICS/ESSIC, University of Maryland, College Park, College Park, MD 20740-3823.
E-mail: jon.mittaz@noaa.gov

DOI: 10.1175/2011JTECHA1517.1

© 2011 American Meteorological Society
thereby a large number of collocated pairs are available, covering a wide range of scene temperatures and zenith angles. Previous results have shown that the current calibration of AVHRR/3 on MetOp-A is introducing scene temperature–dependent biases of up to 0.5 K when compared to IASI (Blumstein et al. 2007; Wang and Cao 2008). Other studies comparing data from the Advanced Along-Track Scanning Radiometer (AATSR), a sensor that is known to be extremely stable and accurate (Smith et al. 2001), support this conclusion (Mittaz and Harris 2008). Long-term studies of the AVHRR calibration have also shown that problems with the AVHRR calibration are by no means limited to the modern AVHRR/3 sensors but are evident in the earliest AVHRRs (e.g., Cao et al. 2006). The fact that the biases seen in the AVHRR are functions of scene temperature and are not simply random implies there is some fundamental error in the AVHRR calibration that needs to be corrected if the AVHRR is to be used for climate studies.

In a recent paper, Mittaz et al. (2009) proposed a new calibration methodology for the AVHRR/3 series of sensors, which, unlike the current operational calibration, is based on a physical approach. This work undertook an extensive analysis of the AVHRR/3 prelaunch data and showed that there were significant problems with the current operational calibration (based on Walton et al. 1998), which was introducing significant biases into the calibrated radiances. By rigorously rederiving the calibration from physical principles, they showed that the origin of these biases was due to a combination of an incorrect parameterization of the fundamental calibration equations coupled with problems with the prelaunch test environment and methodology. In particular, temperature drifts of the instrument during prelaunch testing, together with the effect of scattered or stray light on the calibration, effectively contaminated the prelaunch test data. Mittaz et al. then proposed a completely new calibration that provided a zero-biased and noise-limited solution when compared with the prelaunch data.

What was not clear from the Mittaz et al. (2009) work, however, was how well the new calibration would work in orbit. As was discussed in Mittaz et al. (2009), some of the parameters derived from the prelaunch data will have included residual contamination from the test chamber environment, making it necessary to use an independent top-of-atmosphere (TOA) radiances source to derive an in-orbit calibration. Preliminary work using the AATSR instrument to derive the in-orbit calibration for the AVHRR has already begun and shows promise (see Mittaz and Harris 2008), but such work is complicated by the fact that the AVHRR and AATSR instruments are not collocated, so one must use techniques such as simultaneous nadir overpasses (SNOs; e.g., Cao et al. 2004), or methods such as that proposed by Merchant et al. (2003), to derive TOA radiances to match with those observed by any given AVHRR.

Here we use IASI as our top-of-atmosphere calibration source and derive a new in-orbit calibration for the MetOp-A AVHRR. Because the two instruments are collocated and observe the same field simultaneously, many matchups can be obtained. Further, IASI has the added benefit of having high spectral resolution, which means we can address the issue of possible spectral response function (SRF) shifts.

The paper is organized as follows. Section 2 describes the two instruments—the AVHRR and IASI—and goes into some detail on the calibration issues of the AVHRR instrument. Section 3 details the AVHRR–IASI matchup data used to compare the two instruments. Section 4 describes the result of the comparison and highlights the problems with the current operational AVHRR calibration, as well as a possible solution that can obtain agreement between IASI and AVHRR radiances to $\pm 0.05$ K. Section 4 also discusses the problem of a possible viewing angle bias that can be quite large, at least in cold (cloud-like) temperatures. Section 5 has a discussion of the new AVHRR calibration and section 6 concludes the paper.

2. AVHRR and IASI

a. The calibration of the infrared channels of the AVHRR: Two different approaches

The AVHRR/3 is a broadband imager with three solar channels at 0.63, 0.86, and 1.6 $\mu$m; a mid-IR (solar/thermal) channel at 3.7 $\mu$m; and two longwave channels at 10.8 and 12 $\mu$m, all of which are recorded at 10-bit resolution corresponding to a $\Delta$(Radiance) of 0.0028, 0.18, and 0.21 mW m$^{-2}$ s$^{-1}$ cm$^{-1}$ for the 3.7-, 10.8-, and 12-$\mu$m channels, or a $\Delta T$ of 0.06, 0.28, and 0.27 K at 220 K (where the 3.7-$\mu$m channel exhibits significant digitization resulting from the Planck function) and 0.16, 0.12, and 0.13 K at 290 K. A continuously rotating elliptical scan mirror provides individual scans of the earth together with observations of an “internal” calibration target (ICT; sometimes referred to as the blackbody) and cold space. The latter can be considered a zero external radiance source at thermal infrared wavelengths and is exploited by electronically “clamping” the detector voltage to a fixed-count value of the analog-to-digital converter. This clamping serves both to minimize the effect of the 1/f noise on the observed counts as well as ensure that the 10-bit telemetry covers the maximum possible temperature range. The ICT and space clamp data are then used to generate calibration information for the thermal infrared channels on a scan-by-scan basis.
Here, we consider two calibration methodologies proposed for the thermal infrared channels of the AVHRR: the current operational algorithm proposed by Walton et al. (1998) and also detailed in the National Oceanic and Atmospheric Administration (NOAA)-K, -L, and -M (KLM) user’s guide (Goodrum et al. 2009); and the new physically based method from Mittaz et al. (2009). Both approaches attempt to address the issue of detector nonlinearity for the 11- and 12-μm channels, and both are based on the prelaunch data taken by the manufacturer of the AVHRR: ITT (e.g., ITT 1997). The two methods are, however, significantly different in their approach to the analysis of the prelaunch data. The Walton et al. approach was to fit a model to the counts data recorded when looking at an external calibration target (ECT), and to fit a “negative radiance of space” term together with a nonlinear quadratic model to recover the ECT radiance and hence temperature. Note that in doing this, the Walton et al. calibration completely ignored the ICT data, and all parameters were derived solely from observations of the ECT. In orbit, the final calibration consists of a three-step process. First a “linear” radiance is calculated from the ICT counts and temperature, including a correction for the negative radiance of space. The linear radiance is then used to estimate a correction via a quadratic to compensate for the detector nonlinearity. The correction is then added to the linear radiance term to give the final scene radiance. Fits to the prelaunch data gave a standard deviation to the fit of the order of 0.05 K (J. Sullivan 2008, personal communication).

Mittaz et al. (2009) took a different approach and looked at the AVHRR in a physically consistent manner. This was driven in part by the observation of gain loops in the orbital data of the Walton et al. calibration, but it was also motivated by a desire to understand the implication of a “negative radiance of space,” which at face value is an inherently nonphysical concept. There was also the issue of reconciling the calibration results with the prelaunch observations of the ICT. Weinreb et al. (1990) had previously pointed out that there was a significant mismatch between the measured ICT temperature (measured using four platinum resistance thermistors) and the radiative temperature that was estimated from substituting the ICT count value into the ECT-derived count/radiance equation. This discrepancy could be as large as 0.7 K and is one of the reasons why the prelaunch ICT measurements were completely ignored in the derivation of the Walton et al. calibration. The fact that the radiance source referenced in orbit is ignored when deriving the calibration coefficients from the prelaunch data is likely to have significant consequences for the absolute accuracy of the in-orbit AVHRR thermal IR radiances. Finally, Mittaz et al. also pointed out that the Walton et al. calibration was introducing a variable second-order (nonlinear) term. Not only does this cause problems with the AVHRR level 1b data format (the precision defined for the nonlinear coefficient is not sufficiently high to accurately track variations in the parameter), it also has no physical justification. As noted in Mittaz et al. (2009), the detector temperature of the AVHRR is very stable and so any detector nonlinearity should therefore be virtually invariant.

All of the above issues are discussed and dealt with in detail in Mittaz et al. (2009), who proposed a new calibration algorithm that had a constant nonlinear term for each instrument channel and also had the flexibility to introduce extra sources of radiance into the calibration to deal with issues such as scattered light. The new calibration can also model the instrument self-emission terms explicitly, enabling the calibration to accurately predict the calibration at times when the calibration system is significantly compromised by events such as solar contamination (e.g., Cao et al. 2001). The contaminating radiances seen during the prelaunch testing, however, will have left residual effects on some of the derived calibration parameters, and these effects need to be removed before the calibration can be applied to in-orbit data.

The Mittaz et al. (2009) calibration equation can be written as

\[
R_E = a + \frac{(\epsilon_{ICT} + \rho)R_{ICT} - \alpha' - \gamma(C_S - C_{ICT})^2}{C_S - C_{ICT}} \times (C_S - C_E) + \gamma(C_S - C_E)^2, \tag{1}
\]

where \(C_S\) are the space counts, \(C_E\) are the counts for an earth scene, and \(C_{ICT}\) are the counts seen when looking at the ICT. Here, \(R_{ICT}\) is the radiance of the ICT determined from its temperature; \(\epsilon_{ICT}\) is the emissivity of the ICT, which is estimated theoretically at 0.985 140 (ITT 1997); and \(\gamma\) is a fixed nonlinear term. As noted in Mittaz et al., the terms \(a\), \(\alpha'\), and \(\rho\) are susceptible to contamination/change in the prelaunch testing compared to their in-orbit values. These terms correspond to physical quantities as follows: the term \(a\) is a constant radiance bias term, which is known to be relatively independent of any scattered radiation and is thought to be related to emission from the instrument itself [see Mittaz et al. (2009) and Walton et al. (1998) for a discussion]; the term \(\alpha'\) is a radiance bias term for the ICT observations, which is likely affected by the nonblackness of the ICT; and the term \(\rho\) is related to contamination of the calibration system by a number of different effects—scattered radiance resulting from the nonblackness of the ICT and/or possible corrections to
the deviations of the ICT radiance caused by gradients across the blackbody and/or differences in the true blackbody emissivity relative to the one calculated by ITT. For a fuller description of the derivation of these parameters, see Mittaz et al. (2009). It is important to also note that the characterization of the ICT blackbody was never independently measured and its emissivity was simply a single value calculated by ITT from knowledge of the geometry and a nominal value for the emissivity of the coating (ITT 1997). All of the above parameters have to be refitted based on in-orbit data to get the correct in-orbit AVHRR calibration.

b. IASI

IASI is a completely different type of sensor from AVHRR, consisting of a Michelson interferometer measuring IR spectra between 645 and 2760 cm\(^{-1}\) (corresponding to wavelengths of 3.6 and 15.5 \(\mu\)m) with a spectral resolution of 0.5 cm\(^{-1}\). At nadir the instrument views a region of 3.3\(^\circ\) \(\times\) 3.3\(^\circ\) or 50 km \(\times\) 50 km, and within each field 2 \(\times\) 2 circular pixels (with a 12-km footprint at nadir) are extracted. The total scan angle range is \(\pm 48.3^\circ\), producing 30 total fields of view (FOV) per scan. The calibration to radiance is initially done on board and relies on cold (space) and warm (blackbody) reference targets observed for every scan line. Postcalibration is then done to take into account any temperature and radiative environment effects on the internal blackbody, the variation of the incidence angle reflectivity of the scan mirror, and possible temperature variations of the scan mirror. The estimated accuracy of IASI-derived window channel brightness temperatures is thought to be good. For example, an intercalibration of IASI against the Atmospheric Infrared Sounder (AIRS) has shown the difference between IASI and AIRS to be \(<0.1\) K for nine pseudochannels between 600 and 1500 cm\(^{-1}\) (Blumstein et al. 2007). Comparisons between IASI, AIRS, and the NOAA Real-Time Global Sea Surface Temperature (RTG_SST) product further indicate that IASI is very stable, with a statistically insignificant drift of order 0.02 K over a 6-month period (Aumann and Pagano 2008). In absolute terms, IASI is thought to be within 0.1 K of truth both from intercomparisons with other high-resolution instruments (e.g., Larar et al. 2010), as well as recent comparisons with the AATSIR instrument, based on a very small sample of matches that seem to indicate a slight negative IASI bias in the 11- and 12-\(\mu\)m channels of \(-0.05\) and \(-0.02\) K, respectively (Illingworth et al. 2009), assuming that a correction of 0.2 K is made for the 12-\(\mu\)m channel of the AATSIR (see Smith 2005). The accuracy and stability of IASI, coupled with its high spectral resolution, make it an ideal instrument from which to derive TOA radiances with which to recalibrate the AVHRR.

3. Dataset and method

For the purpose of this study we analyzed both the current operational calibration and the Mittaz et al. (2009) calibration for the AVHRR longwave channels 4 and 5 (at 11 and 12 \(\mu\)m, respectively). Because of spectral coverage issues in IASI at the shortwave end, channel 3B (centered at 3.7 \(\mu\)m) will not be considered. To compare the AVHRR and IASI instruments we have used the IASI spectrum extracted from the IASI level 1c files together with AVHRR data obtained directly from the AVHRR level 1b record. Here we note that using the available AVHRR radiances contained within the IASI level 1c files is not sufficient for our purposes because we wish to recalibrate the AVHRR radiances from the raw AVHRR counts. We therefore need the relevant AVHRR calibration information, including blackbody counts and temperatures, space counts, and scene counts for each AVHRR pixel contained within the IASI FOV.

As was done in Wang and Cao (2008), we first have to exclude inhomogeneous scenes from our analysis. This is because a strong field inhomogeneity impacts the IASI instrument line shape and also makes errors in the collocation of the AVHRR pixels with the IASI field of view much more important. Such errors can be minimized by using homogeneous fields with an estimate of the homogeneity being derived from the AVHRR radiances contained within the IASI FOV. In the study by Wang and Cao (2008) the AVHRR radiances used were those contained in the IASI level 1c files, which consisted of up to six radiance clusters within each FOV, representing different scene types, each of which had a separate mean and standard deviation (Phillips and Schlüssel 2005). Wang and Cao then used the ratio between the weighted average of the standard deviations from the individual clusters and the weighted average of the cluster means (weighted by the appropriate coverage) using a threshold set at 0.01 to find homogeneous fields. While this may seem to be a reasonable test to exclude scene inhomogeneities, it is, in fact, insufficiently stringent. This is because such a test does not take into account the inhomogeneities that are present because of the variation between individual cluster means. Figure 1 shows the histograms of the weighted standard deviations and the estimated standard deviation/variation of the cluster means (which is only calculated when there were at least two clusters), and also shows that the dominant source of variance is, in fact, the differences in the cluster means. Therefore, it is quite likely that
Wang and Cao (2008) included many inhomogeneous fields in their study. Unlike Wang and Cao (2008), however, we have not used the IASI level 1c radiance clusters, but instead have used the AVHRR radiances within each IASI FOV obtained from the AVHRR level 1b files directly. This avoids the problem of combining the inhomogeneity within each IASI radiance cluster with the inhomogeneity of the separate clusters within the IASI FOV to derive a test threshold for exclusion. Deriving a simple mean and standard deviation for the complete IASI footprint, and using a similar ratio to test Wang and Cao’s 0.01 field, now excludes a significant number of fields. Where Wang and Cao excluded around 50% of fields, this new test excludes around 95% of fields and is indicative of the problems with the Wang and Cao test outlined above.

For our analysis we have actually used a different and more consistent homogeneity test than that described above and we have simply removed all of the IASI data, where the standard deviation of the AVHRR pixels contained in each IASI footprint is greater than 1 K, making our test equivalent to selecting footprints containing a single IASI level 1c radiance class. Because there is significantly reduced data sampling because of the more rigorous exclusion test, our samples consist of matches that are taken over a complete day for each date of interest rather than analyzing single orbits of IASI/AVHRR data, as done by Wang and Cao (2008). For each day we extract the AVHRR raw counts and calibration information for each IASI footprint that satisfies our homogeneity test, and we also convolve the AVHRR SRFs for the two window channels with the IASI spectra to generate AVHRR-equivalent radiances and brightness temperatures (BTs). After excluding any extreme outliers at greater than three sigma from the mean difference, we then use either the Walton et al. (1998)-calibrated AVHRR BTs or BTs calculated from the new Mittaz et al. (2009) calibration. This allows the average AVHRR BT over an IASI footprint to be compared with the IASI-estimated AVHRR BTs.

For this study we have extracted 12 daily IASI–AVHRR datasets—two per month for a 6-month period, with the data being taken on the 1st and 15th of the month. For each day that is analyzed there are typically from 5000–6000 matches that are sufficient to determine both the differences between the IASI and AVHRR BTs as well as to derive a new AVHRR calibration. Given the size of the IASI datasets, however, where a 3-min section of IASI level 1c data occupies 60 MB of disk space, we have been limited in the total amount of data it is possible to analyze, and 6 months was chosen as a compromise between studying possible temporal/seasonal variations in the calibration and disk space and processing limitations. The choice of 2 days month$^{-1}$ gives two 6-month-long datasets allowing us some level of independence when we derive a new in-orbit calibration, because we can use one dataset when fitting a new in-orbit calibration (we have used the 15th of the month) and use the second dataset (the 1st of the month) as an independent check on how well the new calibration is working. To do the fitting we have limited ourselves to the data within the 3.3 $\times$ 3.3$^\circ$ nadir view. This is because of a known variability in bias resulting from the satellite zenith angle bias seen in the IASI–AVHRR comparisons that will be discussed further in section 4c.

4. Results

a. Comparing IASI with the Walton et al. calibration

Figure 2 shows the result of comparing 1 day’s worth of matchups for the IASI nadir field of view using the current operational calibration for the longwave IR channels (Walton et al. 1998) to calculate the AVHRR BTs. Both our analysis and that of Wang and Cao (2008) show a scene temperature–dependent bias between IASI and the AVHRR BTs for both the 11- and 12-$\mu$m channels. Also shown in Fig. 2 is the best-fit straight line to the biases, together with the mean and standard deviation derived by binning the biases into 10-K bins, which indicates that the bias trend is, to a first approximation, almost linear. Direct comparison of our results with those of Wang and Cao (2008) shows a similar trend for
the 11-μm channel while the slope for the 12-μm channel from our analysis is somewhat flatter than theirs. The reason for this discrepancy is not clear, although we did not analyze data from the same time period as that of Wang and Cao (our data were taken approximately 1 year later). It is possible that the inclusion of more inhomogeneous fields in the Wang and Cao study, as mentioned above, may have had some channel-dependent impact on their derived trends.

THE BEHAVIOR OF THE WALTON ET AL. BIASES OVER TIME

Because our sample covers a 6-month period, we can also look for long-term trends in the AVHRR–IASI biases. To demonstrate this we have used the bias value at two different temperature ranges [220–230 K (cold) and 290–300 K (warm)] to parameterize the biases. Figure 3 shows these biases plotted as a function of time and illustrates a distinct separation between cold and warm, which is expected because of the scene temperature–dependent bias shown in Fig. 2. However, even though the difference in bias between typical cloudy and clear-sky temperatures is large, the biases are remarkably stable over our 6-month study period, with a typical variation of <0.05 K. Figure 3 actually demonstrates something that has been suspected for some time, which is that the AVHRR itself is an inherently stable instrument.

It is quite likely that, in part at least, the stability of the biases shown in Fig. 3 is primarily due to the constancy of the AVHRR instrument temperature over the same time period. It was shown in Mittaz et al. (2009) that the Walton et al. (1998) calibration suffered from two major bias effects—a trend as a function of scene temperature, similar to that shown in Fig. 2, together with an increasing bias as a function of instrument temperature. Figure 4 shows that the average orbital temperature for the AVHRR on board MetOp-A is very stable and varies by no more than ~0.1 K over our 6-month period of study. Such stability of the average temperature in turn implies that the instrument calibration will also be very stable. However, any relationship between the absolute instrument temperature and the calibration coefficients for earlier AVHRRs, whose instrument...
temperatures could vary significantly over their lifetimes, is less certain and will be the subject of future work.

b. Refitting the Mittaz et al. calibration to the IASI data

As discussed in section 2a, we do not expect that all of the parameters derived for the Mittaz et al. (2009) calibration from the prelaunch data correct for the in-orbit data because of the contaminating sources of radiance present in the prelaunch test environment. To get an in-orbit calibration we must therefore refit the contaminated parameters to derive new values for the $\alpha$, $\alpha'$, and $\rho$ parameters [see Eq. (1)]. As discussed earlier, we expect $\alpha$ to be nonzero because it does not seem to be caused by the prelaunch contaminating radiance but from the radiance of the AVHRR instrument itself; $\alpha'$ should to be close to zero because it will likely have been significantly affected by test chamber radiance via the nonblackness of the ICT; and $\rho$ will have some undetermined value because it can correct for a number of different effects. However, it must be said that the two parameters ($\alpha'$ and $\rho$) are potentially linked in the fitting process and, depending on the quality of the matchup data, may not be uniquely determined. As will be shown, in this case the fitted parameters behave as expected (with the $\alpha'$ term being much smaller than the prelaunch data and close to zero), but the $\alpha'$ and $\rho$ parameters may not be unique. The $\alpha$ parameter as a single bias term, however, is likely to have a unique and physical value. Also note that, in order to ensure a valid starting point for the nonlinear fitting procedure, we have actually fitted an extra radiance term ($\Delta \alpha$) over and above the prelaunch value rather than a new value of $\alpha$ itself.

To fit the data we have used the Interactive Data Language (IDL) package MPFIT (Markwardt 2009).

One other issue regarding the refitting of the calibration has to be addressed before continuing. As was pointed out by Mittaz et al. (2009) some adjustment to the SRF may be required in order to calibrate the AVHRR correctly. In the case of the 3.7-µm channel, Mittaz et al. showed that, without shifting the SRF, the self-emission radiances for many AVHRR instruments were nonphysical. The calibrations for the 11- and 12-µm channels were also improved with SRF shifts. The presence of a shift in the SRF also impacts more than just the conversion from radiance to brightness temperature. Mittaz et al. showed that there was a significant correlation between an SRF shift and the derived instrument nonlinearity, implying that any study of the shift in the AVHRR SRF must take both changes into account. Previous studies of SRF shifts have tended to only concentrate on the radiance-to-temperature conversion part and have not reanalyzed the prelaunch data to derive new nonlinear coefficients (e.g., Wang and Cao 2008). Here we do both.

1) Results

Because of the high degree of stability of the AVHRR (illustrated in Fig. 3), we have used data from all 6 months to fit values for the various calibration parameters. To maintain some level of independence, only the data obtained on the 15th of each month were used in the fitting process. This allows us to compare the IASI and the newly calibrated AVHRR BTs to data taken on the 1st of the month as an independent check. Figure 5 shows the equivalent of Fig. 2 with new fitted parameters (11 µm: $\Delta \alpha = 0.51$, $\alpha' = 0.04$, $\rho = -0.006$; 12 µm: $\Delta \alpha = 0.47$, $\alpha' = -0.005$, $\rho = 0.001$). These new fitted parameters show the behavior that is expected based on the problems noted with the prelaunch data. The overall bias terms remain positive and are, in fact, larger than the prelaunch values, which almost certainly reflect changes in the thermal environment between the prelaunch test chamber and that in orbit. The fit also returns a much smaller ICT bias relative to the prelaunch values ($\alpha'_{11\mu m} = 2.21$ and $\alpha'_{12\mu m} = 1.99$), which again shows the effect of the test chamber contamination on the prelaunch parameters. The differences in the $\rho$ terms between pre- ($\rho_{11\mu m} = 0.0272$ and $\rho_{12\mu m} = 0.0253$) and postlaunch values may in part reflect the fact that the
ITT-calculated emissivity of 0.985 140 was purely theoretical (ITT 1997). However, as mentioned in Mittaz et al. (2009), the \( \rho \) parameter can correct for a number of different effects, with blackbody emissivity being just one. What is true is that both \( \rho \) terms are also much smaller than their prelaunch values, implying that, as with the \( \alpha' \) terms, extra emission from the test environment was a significant contributor to their derived values. Figure 5 shows the new calibration when applied to the independent (1st of the month) data for the 6-month period studied and shows that the new calibration provides a close to zero trend, zero bias solution. Calculating the same biases as described in section 4a gives −0.003 K at 220–230 K and 0.020 K at 290–300 K for the 11-\( \mu \)m channel, and −0.014 K at 220–230 K and 0.012 K at 290–300 K for the 12-\( \mu \)m channel. These biases are significantly reduced compared to the Walton et al. (1998) calibration (approximately −0.11 K at 220–230 K and +0.53 K at 290–300 K for the 11-\( \mu \)m channel and −0.26 K at 220–230 K and +0.31 K at 290–300 K for the 12-\( \mu \)m channel; see Fig. 3).

We can now address the issue of SRF shifts. Figure 6 shows the results from allowing the SRF position to be shifted, where again we have fitted data from the 15th of each month. Note that, as explained earlier, this involves not only fitting new calibration parameters to the IASI–AVHRR dataset but also a reanalysis of the prelaunch data in order to obtain the correct nonlinear coefficient for any given SRF shift. For the 11-\( \mu \)m channel we find that the best-fit SRF shift is +5.3 cm\(^{-1} \), which has made a noticeable improvement to the residual biases. In the case of the 12-\( \mu \)m channel a shift of only +0.29 cm\(^{-1} \) is required with only a slight improvement in the residuals. We note that because the SRF shift is coupled with the nonlinearity and is not independently measured, it is possible that the question of SRF shifts may have to be reinvestigated if one wishes to match radiative transfer modeled radiances with those observed by the AVHRR. Although for the majority of this analysis we use the shifted SRF calibration, the use of unshifted SRFs calibration returns almost the same biases as the shifted returns to typically within one-tenth of a kelvin, which may be adequate for many purposes.

The new calibration reduces the AVHRR–IASI bias to levels close to the expected noise levels of the AVHRR and also reduces the AVHRR–IASI scatter. Figure 7 shows the distribution of the AVHRR–IASI scatter for both the Walton et al. (1998) calibration and the new calibration, together with Gaussian fits to the data. To estimate the intrinsic AVHRR–IASI scatter for the Walton et al. calibration we tried both simple linear and quadratic fits to remove the scene-dependent biases that are inherent in their method. It is apparent that the new calibration reduces the overall scatter compared with both the linear- and quadratic-detrended Walton data, especially for the 11-\( \mu \)m channel. Because the scatter is not exactly Gaussian, we have calculated three different forms of the standard deviation, which are shown in Table 1. While there are differences in the standard deviations resulting from the slightly non-Gaussian nature of the distribution, one thing is clear: the new calibration always reduces the scatter. Compared to the linearly detrended Walton et al. data, the improvement in scatter is (in radiance space) \( \Delta \sigma = 0.033, \Delta \sigma_{\text{Robust}} = 0.039 \) [with the interquartile range (IQR) divided by 1.349, i.e., the relationship between IQR and standard deviation for a normal distribution] and \( \Delta \sigma_{\text{Gaussian}} = 0.038 \text{ mW m}^{-2} \text{ s}^{-1} \text{ cm}^{-1} \) for the 11-\( \mu \)m channel; and \( \Delta \sigma = 0.030, \Delta \sigma_{\text{Robust}} = 0.023, \) and \( \Delta \sigma_{\text{Gaussian}} = 0.024 \text{ mW m}^{-2} \text{ s}^{-1} \text{ cm}^{-1} \) for the 12-\( \mu \)m channel. Relative to the total scatter this then corresponds to a ~20%–30% improvement solely resulting from the improvement in the AVHRR calibration. In terms of a \( \Delta T \) improvement, and using the normal standard deviation as an example, this corresponds to an improvement of 0.022 (11 \( \mu \)m) and 0.019 K (12 \( \mu \)m) at 290 K relative to the total standard deviation of 0.089 and 0.091 K, respectively. Even when using a quadratic model to remove the scene-dependent biases in the Walton et al. case, the new calibration still improves the scatter in both the 11-\( \mu \)m channel (improved by ~20%) and the 12-\( \mu \)m channel, where the improvement is somewhat smaller, around 15% (\( \Delta \sigma = 0.029, \Delta \sigma_{\text{Robust}} = 0.026, \) and \( \Delta \sigma_{\text{Gaussian}} = 0.027 \text{ mW m}^{-2} \text{ s}^{-1} \text{ cm}^{-1} \) for the 11-\( \mu \)m channel, and \( \Delta \sigma = 0.024, \Delta \sigma_{\text{Robust}} = 0.013, \) and \( \Delta \sigma_{\text{Gaussian}} = 0.014 \text{ mW m}^{-2} \text{ s}^{-1} \text{ cm}^{-1} \) for the 12-\( \mu \)m channel.)
channel). Note that the total standard deviations themselves contain contributions from both the IASI and AVHRR noise and thus are an upper limit to the AVHRR noise characteristics; the prelaunch analysis (Mittaz et al. 2009) indicated the AVHRR noise levels were at the 0.036- (11 μm) and 0.027- (12 μm) K level (or 0.055 mW m⁻² s⁻¹ cm⁻¹ and 0.044 mW m⁻² s⁻¹ cm⁻¹ in radiance space).

While the improvement over the Walton et al. (1998) calibration shown in Figs. 5 and 6 is significant, there are residual bias trends in the data with respect to scene temperature. Figure 8 shows the mean bias for data taken from both the unfitted matches (black) and the data to which the calibration has been fitted (gray), binned into 10-K bins. This plot shows essentially the same information as that in Fig. 6, but without the individual matches overlaid, and we have plotted the standard error on the mean rather than the standard deviation. The 12-μm channel shows little evidence for systematic trends in the residuals (apart from an apparent falloff in the bias above 300 K, where the number of matches is reduced), but the 11-μm channel shows some evidence for a

Table 1. Standard deviations for the AVHRR–IASI scatter for the Walton et al. (1998) calibration (after either removing a linear model or quadratic model for the scene temperature–dependent bias) and the new calibration. Listed are the standard deviation, the robust standard deviation, and the Gaussian σ. Units are mW m⁻² s⁻¹ cm⁻¹.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>11 μm (linear)</td>
<td>0.141</td>
<td>0.137</td>
<td>0.090</td>
<td>0.081</td>
<td>0.098</td>
<td>0.090</td>
</tr>
<tr>
<td>12 μm (linear)</td>
<td>0.150</td>
<td>0.147</td>
<td>0.092</td>
<td>0.089</td>
<td>0.101</td>
<td>0.098</td>
</tr>
<tr>
<td>11 μm (quadratic)</td>
<td>0.140</td>
<td>0.137</td>
<td>0.085</td>
<td>0.081</td>
<td>0.094</td>
<td>0.090</td>
</tr>
<tr>
<td>12 μm (quadratic)</td>
<td>0.149</td>
<td>0.147</td>
<td>0.090</td>
<td>0.089</td>
<td>0.099</td>
<td>0.098</td>
</tr>
</tbody>
</table>
systematic trend on both samples with an increasing bias from \( \sim 230-300 \) K of the order of \( 0.06-0.08 \) K. While probably at the limit of what may be reasonably detected, there are a number of possible explanations for such a trend. It may be due to inherent biases within the IASI data (0.05 K is certainly smaller than the expected accuracy of IASI), but, given that both the 11- and 12-\( \mu \)m channels are contained within a single IASI band, a trend in the 11- but not the 12-\( \mu \)m channel would indicate that the 11-\( \mu \)m trend may be an AVHRR rather than an IASI effect. Subtle problems with the 11-\( \mu \)m SRF may explain it, although the lack of an apparent trend in the 12-\( \mu \)m channel SRF makes this unlikely. A more likely explanation is that there are residual problems with our detector model of the AVHRR and, in particular, the quadratic model of the detector (i.e., \( R = \alpha + \beta C + \gamma C^2 \)). That such a model is inadequate has already been shown by Mittaz and Harris (2008) for the particular case of the 11-\( \mu \)m channel for the AVHRR on board NOAA-17, and has also been discussed in Mittaz et al. (2009) in a similar context. While the evidence for the need of a more complex detector model is slight, it will be interesting in the future to look at NOAA-17 in detail to ascertain whether there are truly issues with our simple model of detector nonlinearity.

2) BEHAVIOR OF THE MITTAZ ET AL. BIASES OVER TIME

In section 4b(1) we fitted 6 months’ worth of data and obtained a single recalibration for the AVHRR, but this gives no indication of the temporal behavior of the residual biases. Figure 9 shows the equivalent of Fig. 3 for the new calibration, where the biases at 220–230 and 290–300 K are again plotted as a function of time. The first thing to note is the change in the \( y \)-axis scale from that in Fig. 3, indicating almost an order of magnitude reduction in the range of biases when compared to the Walton et al. (1998) calibration. While the total biases are small, Fig. 9 does seem to show a small drift in the bias in both channels, which, although slight, is almost certainly related to a seasonal variation seen in longer time series (see section 5). Of course, it is also possible that we are approaching the limit of what the AVHRR–IASI data comparison can tell us since we are certainly now at bias levels that are at or below the commonly understood accuracies of both instruments.

c. THE ZENITH ANGLE (SCAN MIRROR) DEPENDENCE

It is already known that there is an apparent zenith angle–dependent bias when comparing the AVHRR with IASI (e.g., Wang and Cao 2008; Blumstein et al. 2007). This effect is also seen in our new AVHRR calibration. Figure 10 shows the bias for the 11- and 12-\( \mu \)m channels for data taken at the first and last IASI swath points at \( 64.8.3.0 \). Unlike the nadir angle data (see Fig. 6) this time there is a strong (up to 1.5 K) scene-dependent bias. This angular dependence can be investigated further. Figure 11 shows the angular dependence for selected sets of scene temperatures, that is, typical cloud temperatures and surface temperatures. This shows that while a zenith angle dependence is apparent at all temperatures, in the temperature regime that is important for determining surface temperatures the effect is tiny (<0.01 K), and thus it can effectively be ignored.
cloudlike temperatures the angular dependence is strong, implying that the effect is related either to the magnitude of the scene radiance or to some property of the emitting source itself, such as polarization. There is also some asymmetry between the left- and right-hand parts of the IASI swath, which has been previously noted by Wang and Cao (2008) and Blumstein et al. (2007) and is thought to be an artifact of the calibration of the IASI instrument. At this point it is not entirely clear what the exact origin of this bias is. A bias resulting from scene radiance can arise due to the $\alpha$ term in Eq. (1) being a function of view angle. Given that we do not know the exact origin of this term with regards to instrument location, it is entirely possible that the amount of scattered light from the instrument body itself changes as the view angle changes. We can, of course, investigate this possibility by refitting the $\alpha$ term for the 48.3° angle case shown in Fig. 10. Figure 12 shows the resulting bias using the refitted $\alpha$ and, while the bias has been reduced, it has not been eliminated and there is still a discernible curve in the residuals. It would therefore seem that a change in the scattered light contribution from the instrument as a function of viewing angle cannot be the sole explanation of the strong angular dependence of the bias, and at least part of the origin of the effect must lie elsewhere.

There are a number of other possibilities to explain this effect. For example, the fault may lie not with the AVHRR but with the IASI data. Because IASI uses a paddle mirror with a varying incident angle, an adjustment has to be applied to correct from the response versus scan angle (RVS) effect, and this may be inadequate. However, the size of the zenith angle bias is probably too large to be explained by errors in the RVS. For example, in terms of broadband infrared sensors, the RVS for the Geostationary Operational Environmental Satellite (GOES) imager in the 12-μm channel can be up to several kelvins (Weinreb, et al. 1997), which is of similar order to the size of the observed bias seen here but, because a correction for RVS is applied during IASI processing, any residual error would be expected to be smaller than this. The simplest explanation may be a degree of polarization sensitivity within the AVHRR instrument, which could help explain the scene temperature dependence because clouds are structured...
entities and are preferentially observed at the colder temperatures. Because the ray path from the target to the detector involves a number of reflections, polarization sensitivity may be introduced at any stage. Furthermore, the radiation observed from cold clouds (i.e., cirrus ice particles, which are structured, e.g., hexagonal columns) may display a polarization signature resulting from preferential scattering angles, even if the particles themselves are randomly organized (e.g., see Takano and Liou 1992), and more so if the particles are preferentially organized, as is frequent in cirrus clouds (e.g., Takano and Liou 1993). The interaction of the partially polarized radiation with the multiple reflections within the instrument could be expected to produce a scan angle–dependent bias, which will appear to be even more prominent when expressed as a brightness temperature resulting from the steepness of the Planck function. However, drawing definitive conclusions on the origin of the observed effect will require a comprehensive study that is beyond the scope of this present work.

5. Discussion

The use of hyperspectral sounders as a cross-platform calibration source is becoming the benchmark for recalibration efforts from a number of groups, including the GSICS project. Previous work, however, seems to have concentrated on either showing the scale of biases seen in the current operational calibration (e.g., Wang and Cao 2008) or providing a simple model for a bias correction after the original calibration has been applied (e.g., Wu et al. 2009). One issue that all such projects have to address is that of defining the matchup criteria to use in such studies, in particular, in reference to any homogeneity tests. For example, Wang and Cao (2008) used what seems to be a fairly lax test on scene homogeneity but showed similar biases to that seen in Fig. 2. We have used a more stringent test, but the question remains as to how important the exact threshold is. To investigate this we have reanalyzed our data with a more stringent homogeneity test wherein we have limited matches to cases where the standard deviation of the AVHRR data within the IASI footprint is only 0.5 K. Figure 13 is the equivalent of Fig. 8 and shows essentially the same bias patterns, implying that as long as the worst cases are excluded, then the exact threshold may have only a small impact on the final result.

In terms of the recalibration methodology, we have also taken a different approach from that of Wang and Cao (2008) or the GSICS project (e.g., Wu et al. 2009), and have used matches between the AVHRR on board the MetOp-A satellite and those of its companion instrument, IASI, to derive a new fundamental AVHRR
calibration by estimating best-fit coefficients of a physical model of the AVHRR instrument response. This new calibration is significantly less biased than the current operational calibration (based on Walton et al. 1998) with biases limited to \(0.05\) K, which is much smaller than the up-to-0.5 K biases seen currently. We have also shown that the in-orbit parameters must be modified from their prelaunch values in part because of the significant contamination of the AVHRR prelaunch data (see Mittaz et al. 2009), but also because of the different thermal environment the AVHRR experiences when in orbit. This demonstrates that, for any accurate AVHRR calibration to be made, two analyses should be undertaken. The first is an analysis of the prelaunch calibration data in order to determine fundamental aspects of the instrument response (e.g., detector nonlinearity), and the second is to modify the calibration parameters for the in-orbit case. Ideally this would involve the use of top-of-atmosphere radiances from well-calibrated sources to tune the calibration parameters, although the calibration may also be corrected on the basis of our understanding of the in-orbit calibration (e.g., with \(\alpha\) tending to zero).

We have also shown that, together with the bias being reduced, the new calibration also reduces the scatter of the AVHRR radiances relative to the IASI radiances, even after empirical bias and trend corrections are applied to the Walton et al. (1998) radiances. Such a finding has implications for the use of satellite-to-satellite intercomparison techniques in correcting calibration errors because often only a simple bias correction is applied to the radiances after the (presumably) incorrect calibration has already been applied. One example of such a correction method is that of the GSICS project, which proposes the use of a simple linear model (fitted using a weighted regression which was not used for our analysis) for at least its geostationary (GEO)–low earth orbit (LEO) bias correction (e.g., Wu et al. 2009). Presumably if such a technique was applied to the AVHRR, the results would be similar to those reported here and the final corrected radiances would be noisier than they have to be. In our case, by going back to the original calibration and correcting the problem at the source one can, perhaps not surprisingly, manage to do better.

Moving on to the relative calibration between IASI and the AVHRR, we have shown that while the 12-\(\mu\)m channel is now close to the limit of obtainable accuracy over much of the observed temperature range (with a typical bias of \(~0.02\) K), there may be some small residual biases in the 11-\(\mu\)m channel at the level of \(0.05\) K. While the origin of such biases is not clear, they may be an indication of residual problems with our model of the AVHRR behavior which, although physically based, is of deliberately minimal complexity. In particular there are two known effects that have not been fully taken into account with our current calibration. First, it is possible that the simple quadratic model of the detector currently used does not completely describe the detector nonlinearity response, and there is supporting evidence for this from the inability of our model to fit the 11-\(\mu\)m channel for NOAA-17 (e.g., Mittaz and Harris 2008). Second, other effects, such as the presence of hysteresis loops in the gain as a function of instrument temperature, need to be studied. Significant loops have been seen in earlier AVHRRs, for example, in NOAA-14 (Trishchenko et al. 2002), and in

---

**FIG. 12.** The AVHRR–IASI bias for data at the edge of the IASI swath (POS = 1) at a scan angle of 48.3° where the \(\alpha\) parameter has been refitted to the data. While the new values of \(\alpha\) have reduced the bias, there is still a residual curve to both the 11- and 12-\(\mu\)m biases, implying that the strong view angle–dependent bias cannot wholly/simply be explained by a change in any sources of radiance from the AVHRR instrument body.

**FIG. 13.** The AVHRR–IASI biases for the 11- and 12-\(\mu\)m channels with a stricter homogeneity test of 0.5 K. The biases show essentially the same behavior as that seen in Fig. 8.
the case of the AVHRR on board MetOp-A they may be the cause of the latitudinal radiance bias corrections required by Merchant et al. (2008). Such loops run contrary to a simplistic expectation of our physical model, which, everything being equal, would predict a monotonic gain change with instrument temperature. The implication is that either some missing time-variable radiance component is contaminating the ICT measurement, or there are significant changes in the thermal gradients across the instrument between day and night. With regard to the latter possibility, it should be noted that although the radiant temperature estimate derived by a simple averaging of the individual ICT PRT measurements is sufficient for a simple linear temperature gradient across the ICT, anything more complex will introduce time-dependent biases into the AVHRR calibration. Work is continuing to resolve these issues.

We have also, for the first time, shown that under the right circumstances, the AVHRR can behave in a very stable manner. While in this work we have concentrated our analysis on a 6-month time scale, we have undertaken a preliminary analysis of a much longer time scale (see Fig. 14). In this plot we have not changed the calibration derived from our original 6-month period. Again, as was evident in Fig. 9, the AVHRR–IASI temperature difference is very stable even over this longer 3-yr period, with an estimated $d\text{BT}$ per decade gradient of $0.055 \pm 0.014$ and $-0.095 \pm 0.011$ for 220–230 and 290–300 K for the 11-µm channel, and $0.060 \pm 0.014$ and $-0.075 \pm 0.010$ for 220–230 and 290–300 K for the 12-µm channel. Such $d\text{BT}$ per decade gradients are slightly larger but still close to those required for climate studies (e.g., Ohring et al. 2005). There also seems to be evidence of a seasonal variation in the temperature difference at the $\pm 0.05$-K level for the cold scene temperatures (220–230 K), which may be caused by subtle changes in the temperature gradients across the AVHRR, or some change in reflected components seen by the calibration system. As was also seen over the original 6-month period, the average instrument (ICT) temperature is very stable (only varying by at a maximum of 0.4 K over 3 years), but also shows seasonal variations that may be linked to the apparent seasonal variation seen in the AVHRR–IASI temperature differences.

Finally, we have verified a significant scan mirror angle dependence of bias between the MetOp-A AVHRR and IASI previously reported by Wang and Cao (2008) and Blumstein et al. (2007), which remains present with our new calibration. Further, we have shown that the nature of this zenith angle bias is a strong function of
scene temperature with the strongest effect (of over 1 K) at high zenith angles occurring only for cold (cloudlike) temperatures. At brightness temperatures used to retrieve surface temperatures the effect is negligible; thus, our new calibration should be able to provide data that are accurate enough to generate at least one of the GCOS-designated essential climate variables, the sea surface temperature (SST). For cloud studies, more work will need to be done to track down the exact nature of the zenith angle bias and its possible relationship to polarization effects.

6. Conclusions

We have shown that, unlike the current operational (Walton et al. 1998) calibration, our new calibration can correct biases and errors for the MetOp-A AVHRR and reduce them to near-negligible levels compared to the IASI instrument. While more work needs to be done on understanding long-term trends in the calibration, especially for the earlier AVHRRs, our new methodology appears to be approaching accuracies close to values consistent with those required for climate change studies (bias <0.1 K and accuracy <0.05 K decade⁻¹), at least for scene radiances used to retrieve surface temperatures. For cloud studies the situation is not so clear cut because there still is a strong zenith angle dependence of the bias. Irrespective, we have shown that by approaching the AVHRR calibration in a physical manner, and taking into account changes in the calibration between prelaunch and postlaunch environments, it is possible to derive climate-ready radiances from the AVHRR, leading to the prospect that some of the essential climate variables, such as the SST, can be derived at the level of stability and accuracy required for climate change studies. It must be stressed, however, that while comparing with the IASI instrument yields small biases, the absolute value of the biases obtained in this work depends on the absolute accuracy of IASI. What this study does show is the power of utilizing hyperspectral sounders as a calibration reference, which will be an invaluable tool for further calibration and climate monitoring activities.

For the future, the methodology of refitting the calibration on the basis of TOA calibration sources will be extended to other AVHRR sensors, in particular the earlier AVHRRs, where the use of accurate TOA sensors such as the (A)ATSR series of instruments (ATSR-1, ATSR-2, and AATSR) can be exploited to extend the data record back to 1991. Comparisons during the (A)ATSR period (1991–present) will also enable us to ascertain differences between the AVHRR/2 and AVHRR/3 instruments. Because our new calibration methodology is physically based, we should be able to extend the calibration prior to 1991 by taking into account the effect of orbit drift and changes in instrument temperature (for which our model accounts) on the AVHRR calibration. The lack of independent TOA radiance data prior to the (A)ATSR series can be overcome to some extent by, for example, utilizing radiative transfer modeling and numerical weather prediction reanalyses. The limitations and error bounds of such an approach can be established by three-way studies using IASI or the (A)ATSR series, where available. We therefore anticipate that the careful application of this new calibration methodology over the complete historic AVHRR data record will eventually yield an accurate fundamental climate data record from the AVHRR for use in at least some climate change studies, bearing in mind the limitations mentioned above.

Acknowledgments. The authors thank Jerry Sullivan, for useful comments and suggestions, and the anonymous referees. This work was partially funded by a NOAA Science Data Stewardship Grant NA09NES4400013.

REFERENCES


